



EVROPSKÁ
UNIE



Katedra softwarového inženýrství, Fakulta informačních technologií,
České vysoké učení technické v Praze

Projekt II - 2

AUDIO RETRIEVAL – KLASIFIKACE HUDEBNÍCH STYLŮ

ZADÁNÍ

Cílem projektu je vytvoření aplikace schopné klasifikovat vstupní nahrávku do jednoho z předem nadefinovaných žánrů. Definice žánrů probíhá tak, že je zvoleno několik klasických skladeb daného žánru, čímž se vytvoří referenční množina skladeb. Při dotazu jsou identifikovány nejpodobnější databázové skladby a na jejich základě je dotaz zařazen do daného žánru.

VSTUP

Audio soubor.

VÝSTUP

Příslušnost vstupního audio souboru do žánru. Vhodné je definovat i míru shody s daným žánrem, případně míru shody se všemi žánry v DB seříděné od nejsilnější shody k nejslabší.

INFORMACE/POTŘEBNÉ ZNALOSTI

Audio signál je popsitelný různými deskriptory. Známa sada deskriptorů, popisující audio z různých pohledů, jsou např. MPEG7 deskriptory. Tento standard popisuje různé deskriptory, které lze z audio extrahovat – např. *AudioWaveform*, *AudioPower*, *AudioSpectrumEnvelop*, *AudioHarmonicity*, ... Každý z deskriptorů popisuje nějakou vlastnost signálu a samotný nepopisuje audio příliš dobře. K dispozici existují nástroje (viz. sekce *knihovny*), které umožňují MPEG7 deskriptory extrahovat. Lze tedy takto relativně levně získat sadu deskriptorů. Tyto deskriptory lze libovolně kombinovat nebo modifikovat v rámci komplexní podobnosti a tak lze získat relativně robustní podobnostní míru na audio.

Každý žánr je vyjadřitelný jako klastř skladeb spadající do daného žánru. Deskriptory a vzdálenost by měly být navrženy tak, aby vzdálenost v rámci skladeb daného klastru byla malá, ale rozdíl mezi jednotlivými klastry by měl být co největší (dobrá separace klastřů).

STAVBA APLIKACE

Aplikace by měla obsahovat části:

- Modul extrakce deskriptorů z audio
- Modul podobnostní míra pro porovnání dvojice audio souborů, tj. jejich deskriptorů
- Modul klastrování.
- Modul identifikace shody audio s žánrem
- Webový interface

POZNÁMKY K ŘEŠENÍ

Obecně je podobnost audio souborů relativně těžký problém. V tomto projektu je třeba navrhnout ne pouze podobnost, nýbrž i zvolit extrahované vlastnosti. Proto audio shoda na úrovni aplikace se shodou, jak ji chápe uživatel, nemusí být příliš dobrá. Jde především o vyzkoušení si návrhu vlastního deskriptoru a podobnosti a vyhodnocení jeho kvality (i nevalná kvalita je vzhledem k netriviálnosti úkolu akceptovatelná ☺).

KNIHOVNY

- *Marsyas* (C++) - <http://marsyas.info/>

- *jAudio* (Java) - <http://sourceforge.net/projects/jaudio/>, <http://jmir.sourceforge.net/jAudio.html>
- *MPEG7 Audio Encoder* (Java) - <http://mpeg7audioenc.sourceforge.net/>

DATA

Aplikace může pracovat pouze s jedním zvoleným datovým typem (midi, wav, mp3, ...). Žánrů by mělo být alespoň 5, každý definován několika skladbami, případně vhodnými úseky skladem (nejvíce odpovídající danému žánru).

EXPERIMENTY

V závislosti na definici extrahovaných vlastností lze testovat vliv jednotlivých parametrů na přesnost a rychlost klasifikace. Parametrizovatelné mohou být jak deskriptory, tak navržená podobnostní míra, tak klastrovací algoritmus, což závisí na konkrétním zvoleném postupu.

ZDROJE

- Přednáška *Feature extraction from audio, music and melody*
- D. Mcennis , C. McKay , I. Fujinaga. JAudio: A feature extraction library, 2005
- J. Foote. An Overview of Audio Information Retrieval, 1998
- J. Foote, M. Cooper, U. Nam. Audio Retrieval by Rhythmic Similarity, 2002
- Cluster Analysis - http://en.wikipedia.org/wiki/Cluster_analysis, 2010